Extended-Range Prediction with Low-Dimensional, Stochastic-Dynamic Models: A Data-driven Approach

Michael Ghil, Mickaël D. Chekroun and Dmitri Kondrashov
Dept. of Atmospheric & Oceanic Sciences and Institute of Geophysics and Planetary Physics,
University of California, Los Angeles, CA
603 Charles E. Young Dr., East, 3845 Slichter Hall
Los Angeles, CA 90095-1567
Ghil phone: (310) 825-1038 fax: (310) 206-5219 email:ghil@atmos.ucla.edu
Chekroun phone: (310) 825-1038 fax: (310) 206-5219 email:mchekroun@atmos.ucla.edu
Kondrashov phone: (310) 825-1038 fax: (310) 206-5219 email:dkondras@atmos.ucla.edu

Michael K. Tippett, Andrew Robertson, Suzana J. Camargo, Mark Cane, Dake Chen,
Alexey Kaplan, Yochanan Kushnir, Adam Sobel, Mingfang Ting, Xiaojun Yuan
Columbia University, Lamont Campus 61 Route 9W,
Palisades New York 10964, USA.
Tippett phone: (845) 680-4420 fax: (845) 680-4866 email:tippett@columbia.edu
Robertson phone: (845) 680-4491 fax: (845) 680-4866 email:awr@columbia.edu
Camargo phone: (845) 365-8640 fax: (845) 365-8157 email:suzana@ldeo.columbia.edu
Cane phone: (845) 365-8344 fax: (845) 365-8157 email:mcane@ldeo.columbia.edu
Chen phone: (845) 365-8496 fax: (845) 365-8157 email:dchen@ldeo.columbia.edu
Kaplan phone: (845) 365-8689 fax: (845) 365-8157 email:alexeyk@ldeo.columbia.edu
Kushnir phone: (845) 365-8669 fax: (845) 365-8157 email:kushnir@ldeo.columbia.edu
Sobel phone: (845) 365-8527 fax: (845) 365-8157 email:ahs129@columbia.edu
Ting phone: (845) 365-8374 fax: (845) 365-8157 email:ting@ldeo.columbia.edu
Yuan phone: (845) 365-8820 fax: (845) 365-8157 email:xyuan@ldeo.columbia.edu

Award Number: N00014-12-1-0911
http://www.onr.navy.mil/sci_tech/32/reports/annual/

LONG-TERM GOALS

The long-term goal of this project is to quantify the extent to which reduced-order models can be used for the description, understanding and prediction of atmospheric, oceanic and sea ice variability on time scales of 1–12 months and beyond.

OBJECTIVES

Demonstrate the ability of linear and nonlinear, stochastic-dynamic models to capture the dominant and most predictable portion of the climate system's variability. Improve the understanding and prediction of the low-frequency modes (LFMs) of variability such as the Madden-Julian Oscillation (MJO), El Niño–Southern Oscillation (ENSO), North Atlantic Oscillation (NAO) and Pacific–North American

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2012	**N/A**	**-**

4. TITLE AND SUBTITLE **Extended-Range Prediction with Low-Dimensional, Stochastic-Dynamic Models: A Data-driven Approach**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Dept. of Atmospheric & Oceanic Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 603 Charles E. Young Dr., East, 3845 Slichter Hall Los Angeles, CA 90095-1567**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**7**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

(PNA) pattern. Validate low-dimensional models (LDMs) based on data sets from observations, reanalyses and high-end simulations.

APPROACH

The major theoretical problem we propose to address is the fact that in a nonlinear system there is a continuum of scales, from the fastest and shortest to the slowest and longest: typically, there is no gap or separation between these scales. Recent ideas and methods from statistical physics and dynamical systems theory — associated with the names of H. Mori and R. Zwanzig — allow one to treat the effect of the larger scales on the smaller ones as memory effects. We propose numerically efficient methods for treating these effects, thus further improving a panoply of methods for model reduction and the application of the reduced models thus obtained to prediction.

The technical approaches include the use of linear and nonlinear, stochastic-dynamic models that capture the dominant and hopefully most predictable portion of the climate system's variability. In particular, the low-order models we propose rely heavily on the proper identification and robust description of LFMs. The expertise of the two teams participating in the project, at UCLA and at Columbia University (CU), includes pioneering work on the use of LFMs in extended-range prediction, as well as extensive experience with the operational performance of a variety of predictive models and methods on several time scales.

Our approach is data-driven and does not require knowledge of the governing equations. Data — whether observations from nature or GCM output — are used to estimate both the model's low-order, deterministic part and its driving noise. These LDM models can be linear (Penland, 1989) or nonlinear, as in the recently developed *empirical model reduction* (EMR) methodology (Kondrashov et al., 2006, 2011, 2005; Kravtsov et al., 2005, 2009). EMR models account for nonlinearity, serial correlation in the noise residue and memory effects, and was shown to be particularly effective when there is no *spectral gap* (time scale separation) between slow and fast variables.

A key advantage of the data-driven approach is that there are no explicit assumptions on the full dynamical equations, and, in fact, the equations do not have to be known. The main difficulty in this approach, though, is that the more realistic an LDM, the more parameters have to be estimated from the limited data; still, when applying well-established regularization techniques, robust parameter estimates can be obtained.

Relatively simple data-driven linear LDMs had already success simulating and predicting a wide range of climate phenomena including ENSO, tropical Atlantic sea surface temperatures (SSTs), storm track variability and the Madden-Julian Oscillation (MJO) (Jiang et al., 2008; Penland and Matrosova, 1998; Winkler et al., 2001). Barnston et al. (2012) analyzed the real-time forecast performance of EMR-ENSO model (Kondrashov et al., 2005) in IRI's ENSO multi-model prediction plume for the 2002–2011 interval, and have found that it "has the highest seasonally combined correlation skill among the statistical models, exceeded by only a few dynamical models".

By extending the EMR results of Kondrashov et al. (2005) beyond the sampled-PDF approach, Chekroun et al. (2011) have developed a pathwise *past noise forecasting* (PNF) method of prediction for nonlinear stochastic systems that exhibit LFMs. The PNF method exploits information on the estimated noise path on which the inverse stochastic model lives, and the LFMs resulting from the interaction between the model's deterministic, nonlinear components and the stochastic ones. The PNF method can

provide retrospective ENSO forecasts at 14–16 months with an accuracy that is comparable to the one obtained at 6–8 months by the methods — dynamical as well as statistical — that are currently available in the IRI plume. We plan to extend this PNF method to other applications and LDMs.

Theoretical track. Provide theoretical guidance for practical development of next-generation LDMs for climate modeling and prediction by developing further and applying the Mori–Zwanzig (MZ) formalism. In the MZ approach, one considers the full nonlinear dynamics as a sum of three terms: (i) the projected dynamics that governs the resolved variables, (ii) the orthogonal dynamics of the unresolved ones, and (iii) their interactions through memory terms that arise in an integral form; the latter are given by repeated convolutions between decaying memory kernels and the resolved modes (Chorin et al., 2002, 2006; Chorin and Stinis, 2006; Grabert, 1982; Stinis, 2006). These integral terms appear in the projected dynamics in order to represent the interaction terms between the resolved and unresolved scales, without assuming a gap or separation between these scales.

Michael Ghil, Mickaël D. Chekroun, Dmitri Kondrashov.

Methodological improvements. Improve further the EMR and PNF methodology to develop computationally efficient and robust LDMs. We plan to develop the new-generation EMR models in which the generalized energy is conserved. This development will follow the spectral expansion of the governing equations in most fluid-dynamical problems, which imposes additional properties on the model coefficients (Ghil and Childress, 1987). We are in the process of introducing appropriate linear constraints in the least-square estimation of the EMR coefficients, thus reducing effectively their number and achieving therewith a form of regularization. The choice of basis in which the resolved variables are represented will also be investigated accordingly, since the typical choice of the EOFs as basis functions for the resolved dynamics is not supported by theoretical guidance.

Michael Ghil, Mickaël D. Chekroun, Dmitri Kondrashov.

ENSO and coupled variability on seasonal time-scales. Improve understanding and prediction of ENSO, tropical Atlantic variability (TAV), and the Indo-Pacific coupled mode (IPT).
Mark A. Cane, Dake Chen, Alexey Kaplan.

Sea ice. Predict Arctic sea ice on intraseasonal to seasonal time scales, improve Antarctic sea ice forecasts and assess the impact of sea ice on mid-latitude predictability.
Xiaojun Yuan, Dake Chen.

MJO. Incorporate slow moisture dynamics into empirical prediction models. Use observations to improve low-order dynamical MJO models.
Adam Sobel, Daehyun Kim.

Extratropical variability and predictability. Determine the extent to which extratropical monthly and seasonal LFV — i.e., PNA, NAO, as well as other regional blocking patterns — can be skillfully predicted from one to 12 months in advance in a LDM including ENSO, MJO and stratospheric linkages. Provide a statistical characterization of extratropical storms and extremes and link these to LFV modes.
Mingfang Ting, Yochanan Kushnir, Andrew W. Robertson.

Downscaling weather and seasonal climate. Use LDMs to relate near-surface temperature and precipitation with predictable climate forcing such as the SST field. Explore the development of statistical tropical cyclone forecasts on the intraseasonal time scale, based on LDM forecasts of ENSO

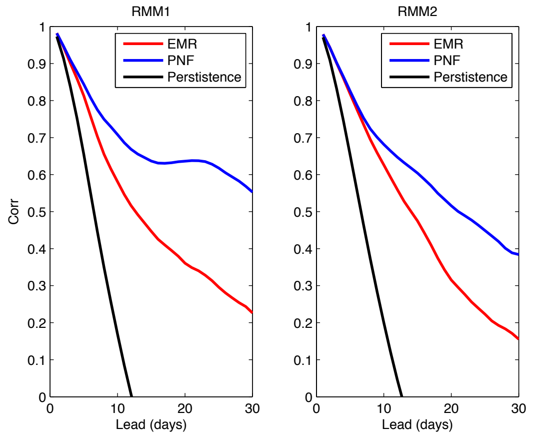

***Figure 1:** Prediction skill for the (RMM1, RMM2) pair of EOF-filtered indices[3] of the Madden-Julian Oscillation (MJO): EMR (red) and PNF (blue) prediction, while the black curve shows damped persistence for a base comparison.*

and the MJO. Target tropical monsoon climates, with particular attention to environmentally and societally relevant sub-seasonal characteristics of weather. The characteristics of interest include the start and end dates of the monsoon seasons, as well as extremes of rainfall and drought within the monsoon season.

Andrew W. Robertson, Suzana J. Camargo, Yochanan Kushnir, Michael K. Tippett.

WORK COMPLETED

Year 1 plans: Complete preliminary study of EMR-PNF prediction for the MJO, and initiate development of energy-conserving EMR framework at UCLA. Interact closely with the corresponding MJO and sea ice sub-teams at Columbia University (CU) to initiate application of these methods to sea ice variability and prediction.

Kick-off meeting planned for October 5 in Palisades, NY, to bring together UCLA and CU researchers with ONR program managers.

Post-doctoral research positions at CU have been advertised, and the post-doctoral position at UCLA will be advertised as soon as financial arrangements have been completed.

RESULTS

- The constructed analogue (CA) forecast method gives weights to past states according to the extent to which they represent analogues of a future state (van den Dool, 1994, 2006). CA is used in operational seasonal forecasting and had been described as having properties that are distinct from those of linear regression. We have shown that standard implementations of the CA method correspond to principal component regression and ridge regression (Tippett and DelSole, 2012).
- A quadratic, three-level EMR model was developed to model and predict the leading pair (RMM1, RMM2) of real-time, daily MJO indices. The EMR model was fitted, and the PNF method was trained, on the first 30 yr of the Real-time Multivariate MJO Index (RMM) record; the respective forecasts were validated on the record's last 3.5 yr. The cross-validated prediction skill of the PNF method is notably better than the EMR alone (blue vs. red curve in Fig. 1) and compares favorably with both the statistical and dynamical results of Kang and Kim (2010).

[3] http://cawcr.gov.au/staff/mwheeler/maproom/RMM/

- As the Madden-Julian oscillation (MJO) moves eastward from the Indian to the Pacific ocean, it typically accelerates, becomes less strongly coupled to convection, and becomes more similar to a dry Kelvin wave. This transition was analyzed using observations of outgoing longwave radiation and ERA Interim reanalyses of surface pressure and 850 hPa zonal wind. Transitions are well defined, with distinct disturbances on either side of the transition whose identities as MJO or Kelvin waves are clear. Transitions occur over a wide range of longitudes in different events, all the way from the Eastern Indian to the Central Pacific oceans.

  We had shown in a simple linear model of intraseasonal moisture modes (Sobel and Maloney, 2012) that wind-evaporation feedbacks induce westward propagation in an eastward mean low-level flow. We consider here additional processes that provide effective sources of moist static energy to the disturbances, and which also depend on the low-level wind. Several processes can act as positive sources in perturbation easterlies: zonal advection (if the mean zonal moisture gradient is eastward), modulation of synoptic eddy drying by the MJO-scale wind perturbations, and frictional convergence. If the sum of these is stronger than the wind-evaporation feedback — as observations suggest may be the case, though with considerable uncertainty — the model produces unstable modes that propagate weakly eastward, relative to the mean flow.

IMPACT/APPLICATIONS

RELATED PROJECTS

REFERENCES

Barnston, A. G., M. K. Tippett, M. L. Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002-2011 — is our capability improving? *Bull. Am. Meteor. Soc.*, **93(5)**, 631–651, URL http://dx.doi.org/10.1175/BAMS-D-11-00111.1.

Chekroun, M., D. Kondrashov, and M. Ghil, 2011: Predicting stochastic systems by noise sampling, and application to the El Niño-Southern Oscillation. *Proc. Natl. Acad. Sci. USA*, **108**, 11 766–11 771, doi:10.1073/pnas.101575.

Chorin, A., O. Hald, and R. Kupferman, 2002: Optimal prediction with memory. *Phys. D*, **166**, 239–257.

Chorin, A. J., O. H. Hald, and R. Kupferman, 2006: Prediction from partial data, renormalization, and averaging. *J. Sci. Comput.*, **28 (2-3)**, 245–261, doi:10.1007/s10915-006-9089-5.

Chorin, A. J. and P. Stinis, 2006: Problem reduction, renormalization, and memory. *Commun. Appl. Math. Comput. Sci.*, **1**, 1–27.

Ghil, M. and S. Childress, 1987: *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory and Climate Dynamics*. Springer-Verlag, New York/Berlin/London/Paris/ Tokyo.

Grabert, H., 1982: *Projection Operator Techniques in Nonequilibrium Statistical Mechanics*. Springer-Verlag, Berlin, x+164 pp.

Jiang, X., D. E. Waliser, M. C. Wheeler, C. Jones, M.-I. Lee, and S. D. Schubert, 2008: Assessing the skill of an all-season statistical forecast model for the Madden–Julian oscillation. *Mon. Wea. Rev.*, **136**, 1940–1956, doi:10.1175/2007MWR2305.1.

Kang, I.-S. and H.-M. Kim, 2010: Assessment of MJO predictability for boreal winter with various statistical and dynamical models. *J. Climate*, **23**, 2368–2378.

Kondrashov, D., S. Kravtsov, and M. Ghil, 2006: Empirical mode reduction in a model of extratropical low-frequency variability. *J. Atmos. Sci.*, **63**, 1859–1877.

Kondrashov, D., S. Kravtsov, and M. Ghil, 2011: Signatures of nonlinear dynamics in an idealized atmospheric model. *J. Atmos. Sci.*, **68**, 3–12.

Kondrashov, D., S. Kravtsov, A. W. Robertson, and M. Ghil, 2005: A hierarchy of data-based ENSO models. *J. Climate*, **18**, 4425–4444.

Kravtsov, S., D. Kondrashov, and M. Ghil, 2005: Multilevel regression modeling of nonlinear processes: Derivation and applications to climatic variability. *J. Climate*, **18**, 4404–4424.

Kravtsov, S., D. Kondrashov, and M. Ghil, 2009: Empirical model reduction and the modeling hierarchy in climate dynamics. *Stochastic Physics and Climate Modeling*, T. N. Palmer and P. Williams, Eds., Cambridge Univ. Press, 35–72.

Penland, C., 1989: Random forcing and forecasting using principal oscillation pattern analysis. *Mon. Wea. Rev.*, **117**, 2165–2185.

Penland, C. and L. Matrosova, 1998: Prediction of Tropical Atlantic Sea Surface Temperatures Using Linear Inverse Modeling. *J. Climate*, **11**, 483–496.

Sobel, A. H. and E. D. Maloney, 2012: An idealized semi-empirical framework for modeling the Madden-Julian oscillation. *J. Atmos. Sci.*, **69**, 1691–1705.

Stinis, P., 2006: A comparative study of two stochastic mode reduction methods. *Phys. D*, **213 (2)**, 197–213, doi:10.1016/j.physd.2005.11.010.

Tippett, M. K. and T. DelSole, 2012: Constructed analogues and linear regression. *Mon. Wea. Rev.*, submitted.

van den Dool, H., 2006: *Empirical Methods in Short-Term Climate Prediction*. Oxford University Press, 240p.

van den Dool, H. M., 1994: Searching for analogues, how long must we wait? *Tellus A*, **46**, 314–324.

Winkler, C. R., M. Newman, and P. D. Sardeshmukh, 2001: A Linear Model of Wintertime Low-Frequency Variability. Part I: Formulation and Forecast Skill. *J. Climate*, **14**, 4474–4494.

PUBLICATIONS

Ghil, M., 2012: Mathematical theory of climate sensitivity, in *Climate Change: Multidecadal and Beyond*, C. P. Chang, M. Ghil, M. Latif and J. M. Wallace (Eds.), World Scientific Publ. Co./Imperial College Press, to appear.

Sobel, A., and E. Maloney, 2012: Moisture modes and the eastward propagation of the MJO. *J. Climate*, submitted.

Sobel, A. H., and D. Kim, 2012: The MJO-Kelvin wave transition. *Geophys. Res. Lett.*, submitted.

Tippett, M. K., and T. DelSole, 2012: Constructed analogues and linear regression. *Mon. Wea. Rev.*, submitted.